Cell

Building Enhancers from the Ground Up: A Synthetic Biology Approach

Roee Amit,[1,3,4,*] Hernan G. Garcia,[2] Rob Phillips,[1,3] and Scott E. Fraser[1,3]
[1]Division of Biology
[2]Department of Physics
[3]Division of Engineering and Applied Science
California Institute of Technology, Pasadena, CA 91125, USA
[4]Present address: Department of Biotechnology and Food Engineering, Technion, Haifa 32000, Israel
*Correspondence: roeeamit@technion.ac.il
DOI 10.1016/j.cell.2011.06.024

SUMMARY

A challenge of the synthetic biology approach is to use our understanding of a system to recreate a biological function with specific properties. We have applied this framework to bacterial enhancers, combining a driver, transcription factor binding sites, and a poised polymerase to create synthetic modular enhancers. Our findings suggest that enhancer-based transcriptional control depends critically and quantitatively on DNA looping, leading to complex regulatory effects when the enhancer cassettes contain additional transcription factor binding sites for TetR, a bacterial transcription factor. We show through a systematic interplay of experiment and thermodynamic modeling that the level of gene expression can be modulated to convert a variable inducer concentration input into discrete or step-like output expression levels. Finally, using a different DNA-binding protein (TraR), we show that the regulatory output is not a particular feature of the specific DNA-binding protein used for the enhancer but a general property of synthetic bacterial enhancers.

INTRODUCTION

A classic view of transcriptional regulation in bacteria is built around the idea of regulated recruitment of RNA polymerase and the dissociable sigma factor σ^{70}. In this picture, the presence or absence of RNA polymerase at a promoter of interest is dictated by the corresponding presence or absence of batteries of transcription factors that either increase (activators) or decrease (repressors) the probability of polymerase binding. An increasingly sophisticated understanding of this kind of regulatory response has resulted in an explosion of efforts in synthetic and systems biology built using a broad palette of different activators and repressors for a range of different promoters of this kind (Bintu et al., 2005b; Elowitz and Leibler, 2000; Gardner et al., 2000; Joung et al., 1993; Müller et al., 1996; Mukherji and van Oudenaarden, 2009 and references therein).

Another whole set of bacterial promoters utilize an alternative sigma factor (σ^{54}) that, together with RNAP, forms a stable closed promoter complex that, unlike its σ^{70} counterpart, is unable to initiate transcription by itself (Buck et al., 2000; Rappas et al., 2007). This effectively causes the polymerase to be poised at the gene of interest, awaiting the arrival of a transcription factor partner that we term the "driver," which releases the polymerase. Consequently, these promoters are regulated in a different fashion than their recruitment counterparts. The activating or transcription driving complex is typically widely separated from the promoter (100–1000 bp) (Ninfa et al., 1987), precluding it from forming direct contact with the poised polymerase. It has been asserted (Huo et al., 2006; Schulz et al., 2000; Su et al., 1990) that DNA looping and ATP hydrolysis are required to induce open complex formation and transcription initiation (Rappas et al., 2007). These regulatory regions belong to a different class of regulatory elements called enhancers, which are more commonly associated with eukaryotic organisms. On its own, a poised promoter has the capability to execute little or no transcriptional regulation, but together with enhancers, they can express their full regulatory potential (Davidson, 2001; Magasanik, 1993).

Enhancer elements are ubiquitous in genomes from all domains of life (Buck et al., 2000; Ninfa and Atkinson, 2000; Rappas et al., 2007). It is hypothesized that enhancers execute their regulatory program by making direct contact with the basal promoter via DNA or chromatin looping. In general, they are made up of contiguous genomic regions that stretch from tens to thousands of base pairs and contain several binding sites for a variety of transcription factors (TF); often, their regulatory output is independent of their location or orientation relative to the basal promoter (Driever et al., 1989; Huo et al., 2006; Ninfa et al., 1987). As a result, enhancers, like gene regulatory networks themselves, can be viewed qualitatively (Davidson, 2006) as modular genomic entities made of three connected irreducible parts: the driver-binding sites responsible for initiation of transcription, protein-binding sites within the enhancer that are responsible for the regulation or modulation of expression levels, and the poised promoter. Whereas other aspects of gene regulation are becoming better defined (e.g., the input/output relationship between different genes in gene regulatory networks) (Bintu et al., 2005a; Kuhlman et al., 2007; Rosenfeld et al., 2005; Garcia

and Phillips, 2011), the underlying mechanisms of regulatory "action at a distance" that are responsible for integrating the various inputs in enhancers remain poorly understood.

To explore the kinds of action at a distance mechanisms that can yield complex regulatory behavior associated with enhancers, we opted to construct synthetic enhancers de novo. In this case, the synthetic approach permits us to systematically construct enhancers in a modular fashion, starting with a minimal enhancer made of driver-binding sites and the poised promoter region and progressively increasing the synthetic enhancer's complexity with the addition of discrete sets of defined enhancer-binding protein-binding sites (TetR or TraR in our case) that are not thought to interact directly with either the driver protein or the poised RNA polymerase. The synthetic approach provides us with an experimental foundation that can be utilized to develop thermodynamic models in which the various states of occupancy of the promoter and their associated statistical weights can be computed and used to explore the enhancer's regulatory output.

We hypothesized that a rich interplay between experiment and theory would not only allow us to increase our predictive capability with respect to enhancer regulatory output, but also tease out the underlying mechanisms for regulatory action at a distance by ensuring that the model and experiment be consistent at every stage of the cascade. At each experimental stage, when an increasingly complex set of regulatory architectures was characterized, the starting point for the theoretical description was the model utilized to describe the more simplified constructs explored during the previous step. Thus, throughout the paper, we will repeatedly resort to thermodynamic models, which exploit equilibrium statistical mechanics to serve as a conceptual framework for all of the experiments.

RESULTS

Expression Levels Are Controlled by DNA Looping

We selected the bacterial NRI/NRII (NtrC/NtrB) two-component system (Magasanik, 1993), controlling nitrogen assimilation in many prokaryotes, to test our methodology. We constructed minimal enhancers using driver-binding sites for the phosphorylated DNA-binding isoform of NRI (NRI∼P) and coupled them to a poised σ^{54} promoter with a DNA linker of varying length. The dimeric NRI∼P proteins assemble on the DNA to form a hexameric complex, which in turn functions as the transcriptional driver in our system. An mCherry reporter was used to measure the transcriptional activity of this promoter (for circuit details, see "Theory: Model for Looping Initiated Transcription" in the Supplemental Information available online).

We reasoned that systematically varying the length of the DNA sequence between the driver-binding sites and σ^{54} promoter will yield an expression pattern that depends on the length of the looped DNA and on the phasing of the complex (the orientation of the driver with respect to the polymerase bound to the promoter that depends on the DNA helical periodicity) in much the same way that phasing impacts expression levels in different looping regulatory contexts (Law et al., 1993; Lee and Schleif, 1989; Müller et al., 1996). In order to check the validity of this assumption, we cloned into the spacer region of the synthetic enhancer 65 distinct DNA sequences (Table S1 and Table S2) of variable length (28–315 bp; Figure 1A and Figure S1A). We carried out fluorescence measurements in bulk while the strains were growing in midlog phase and subsequently normalized the fluorescence levels obtained for each strain to the value measured for the maximally expressing strain (L = 70 bp).

At first glance, the results shown in Figure 1B seem to exhibit a strongly fluctuating behavior with a nontrivial dependence on looping length (L). However, a useful framework for considering this complex data is provided by the thermodynamic model schematized in Figure 1C, which invokes a model inspired by the underlying DNA biophysics of looping, transcriptional mechanics, and equilibrium thermodynamics (see "Theory: Model for Looping Initiated Transcription" in the Supplemental Information, Figure S2, and Figure S3). The essence of the model depicted in the figure is that there are two states of interest, both of which have the $(NRI\sim P)^6$ hexamer and RNA polymerase (RNAP) bound but only one of which is looped and transcriptionally active. The looped state is weighed by a looping J factor (a measure for the local concentration of the hexamer in the vicinity of RNAP) and a dissociation constant between the $(NRI\sim P)^6$ hexamer and RNA polymerase. To simplify the interpretation of the results, we collapse the looping J factor and the dissociation constant by defining the ratio J/K_{nr} as the looping capacity $\chi(L)$.

The model generates a fit that rises rapidly for L < 70, slowly declines for L > 70 (light blue dashed lined), and is modulated by a characteristic periodicity of 11.0 ± 0.1 bp. This value for the periodicity likely corresponds to the helical period of the DNA itself and is in good agreement with previous measurements (Becker et al., 2005; Law et al., 1993; Lee and Schleif, 1989; Müller et al., 1996). It is worth noting that, whereas the error to the fit of the periodicity exhibited by our data is low, the rest of the parameters, which characterize the looping capacity function, cannot be determined to a high level of certainty. As shown in Figures S2A and S2B, various candidates for the looping capacity function can generate plausible envelope functions for the data, as shown by the red line in Figure 1B and Figure S2B. Discriminating between alternative looping capacity functions would require data from larger loop lengths than those obtained here.

Enhancer Repression Is a Bimodal Function of Spacer Length

Given that the level of transcription depends critically upon DNA looping, we reasoned that, by installing binding sites for other transcription factors within the looped region, we might tune the propensity for loop formation and hence the level of expression by controlling the concentration of the active transcription factors. We suspected that one possible way of generating this effect was by making the intervening DNA more rigid though the binding of a common repressor TetR, whose binding to DNA is thought not to induce long-range deformations (Ramos et al., 2005 and references therein). This, in turn, would lead to an inhibition of the looping process, which would result in the repression of the synthetic enhancer circuit, yielding a reduction in the quantity of the fluorescent reporter.

In order to test this assertion, we added cassettes to the synthetic enhancer containing one, two, three, or six binding sites for TetR. The cassettes were cloned 28 bp downstream of the

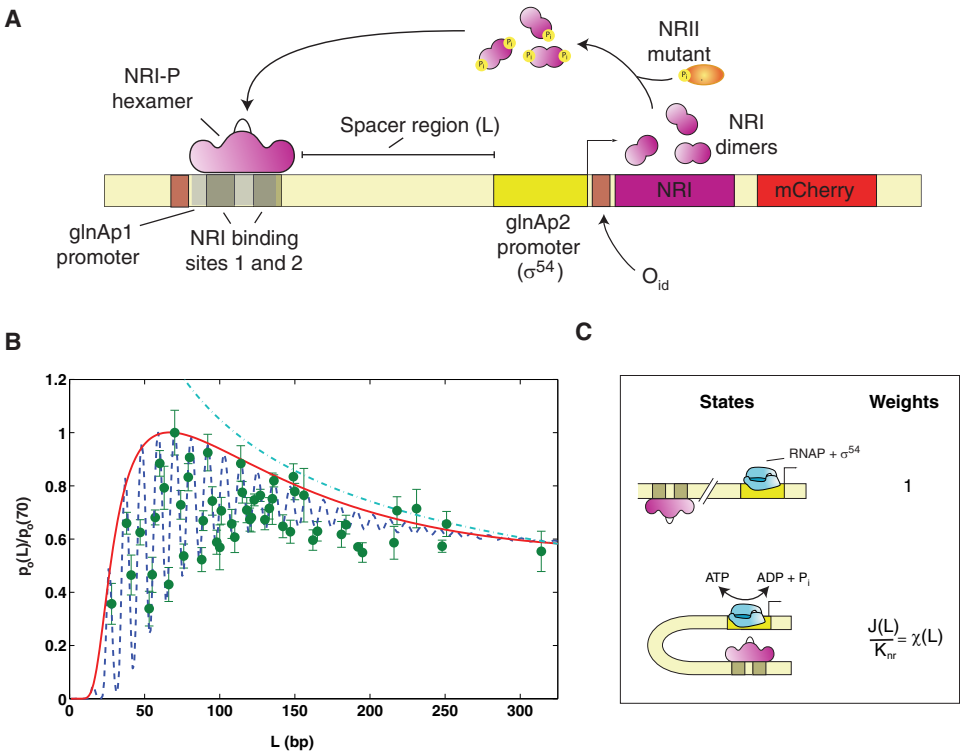

Figure 1. Enhancer Activation Depends Strongly on Looping

(A) Schematic for synthetic enhancer circuit. In short, the circuit expresses via a σ^{54} promoter the *glnG* (*ntrC*) gene, whose protein product (NRI) remains phosphorylated at all times via the action of the phosphatase-deficient mutant NRII2302 (Atkinson et al., 2003), which also serves to decouple the NRI/NRII system from the nitrogen assimilation pathway. The synthetic enhancer circuit was transformed into a *ΔGlnL:ΔGlnG:3.300 E. coli* strain (3.300LG) on a low-copy plasmid (≅10/cell).

(B) Relative fluorescence level $p_o(L)/p_o(70)$ versus looping length L data (green circles). For each looping length, $p_o(L)/p_o(70)$ is defined as the ratio between the measured fluorescence level of the synthetic enhancer strain to the fluorescence level of the brightest strain (L = 70 bp, the natural glnAp2 enhancer looping length). The fits correspond to our expression model with (blue dashed line) and without (red line) the periodic modulation (see "Theory: Model for Looping Initiated Transcription" in the Supplemental Information for more details). The light-blue dashed line corresponds to a fit by an empirical power-law decay curve of power −1/2. Error bars correspond to the standard deviation from multiple measurements.

(C) Schematic Model for enhancer-activated transcription for our constructs, which requires ATP hydrolysis and DNA looping to bring the driver/activator protein complex into physical contact with the "poised" σ^{54}-RNAP complex.

See also Figure S1, Figure S2, and Figure S3.

NRI#2 binding site (Figure 2A) to ensure that the first TetR (Hillen et al., 1984) does not interfere with the binding of the NRI∼P complex (Hervás et al., 2009). This isolates the repression effects to a modification of the looping capacity function when TetR is present, the description of which is developed in "Theory: Model for Looping Initiated Transcription" in the Supplemental Information. The extent of repression for each cassette was quantified by measuring the fluorescence of the reporter both in the presence of a high number of TetR proteins and in their absence. In Figure 2B, we plot repression values as a ratio of the repressed to the unrepressed fluorescence levels for each synthetic enhancer circuit as a function of the looping DNA length (as defined in Figure 2A). The figure shows the experimental data for the 1-Tet (one TetR-binding site), 2-Tet, and 3-Tet cassettes. For all cassettes used in the experiment, the data show a signature for bimodality with either strong repression for synthetic enhancer lengths $L < L_t$ or weak repression for lengths $L > L_t$. The length L_t, which serves as a DNA length scale setting a sharp transition between the two repression regimes, varies for each cassette type (labeled as L_{t1}, L_{t2}, and L_{t3} on the plots) and seems to depend systematically on the number of binding sites and the size of the binding region of TetR (Hillen et al., 1984).

In order to understand the bimodal behavior, it is instructive to consider the short and long loop length limits. For short loop lengths, one simple interpretation is that the DNA-TetR complex behaves like a "rigid" nucleoprotein complex with an effective persistence length longer than that of bare DNA. Alternatively, the heightened repression at short looping lengths could be due to some other biophysical mechanism that promotes TetR-induced interference with the ability of the NRI∼P complex to loop. Either way, for $L < L_t$, looping is far less likely to take place, and the RNAP will remain poised.

For long loop lengths, wherein the weak repression regime is observed, the rigidification effect hypothesized for lower lengths is diminished. In this regime, the data indicate that repression levels are weakly dependent on the loop length and the synthetic

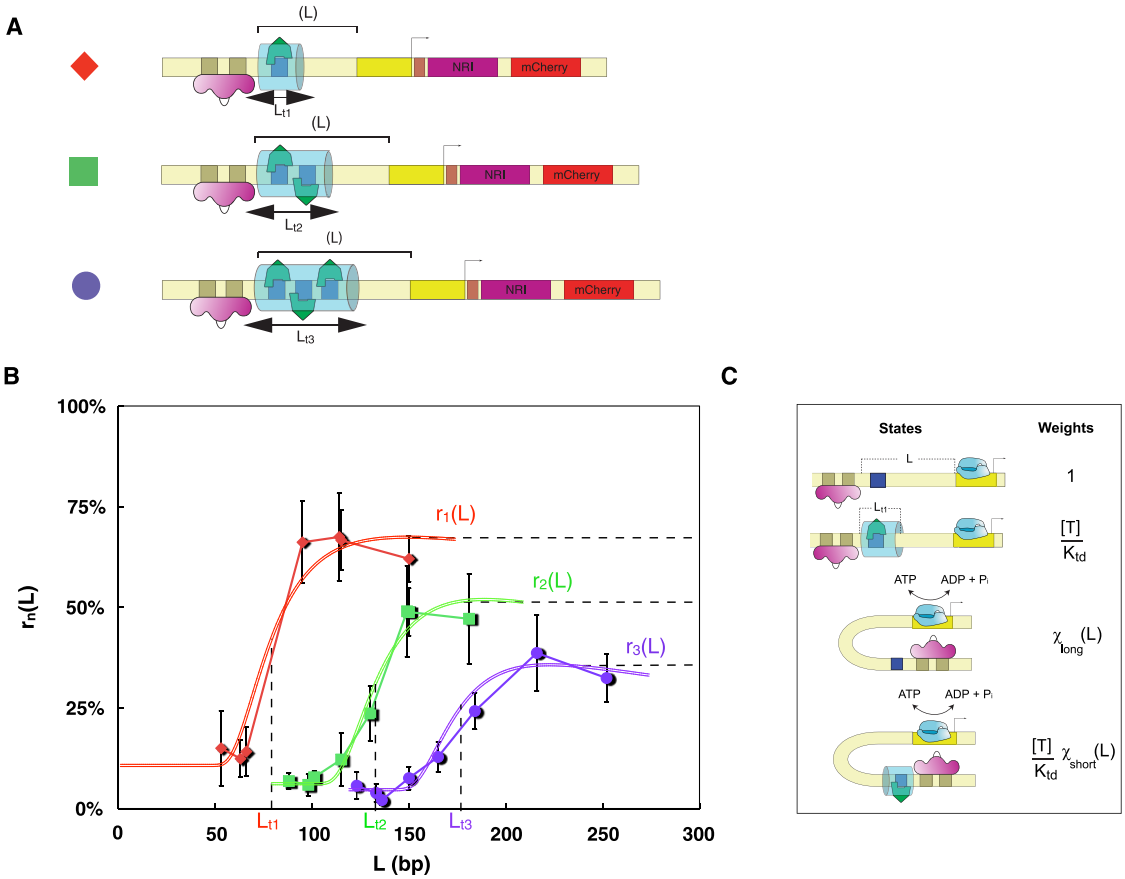

Figure 2. Bimodal Repression of Enhancer-Based Transcription

(A) Schematic showing the constructs used to study enhancer repression containing 1-, 2-, and 3-TetR-binding sites, respectively. The binding sites for TetR are positioned 28 bp upstream of the NRI#2 site and with 16 bp spacing for the 2- and 3-Tet cassettes. The TetR-rigidified region of the spacer DNA (denoted by light-blue shade and L_{t1}, L_{t2}, and L_{t3}) is hypothesized to be the mechanism responsible for repression.

(B) Expression data exhibiting bimodal behavior for the 1-Tet (red diamonds), 2-Tet (green squares), and 3-Tet (purple circles) cassettes. The data are depicted as percent relative to the unrepressed expression levels for the 1-, 2-, and 3-Tet cassettes, respectively. r_1(L), r_2(L), and r_3(L) levels correspond to the repression functions as defined in "Theory: Model for Looping Initiated Transcription" in the Supplemental Information (Equations S27–S29). The values of these functions at particular lengths are used as input to the model and fits (for the data shown in Figure 3, Figure 4, Figure 6, and Figure S5). The colored curved double lines for each data set correspond to model fits (see Figure S4 for additional detail), and the dashed lines correspond to the length-independent repression value that each cassette seems to approach. Error bars correspond to the standard deviation from multiple measurements.

(C) States and weights schematic for the model used to describe the 1-Tet repression data. The two additional states correspond to the looped and unlooped configurations of the DNA with TetR bound to the enhancer.

See also Figure S4.

enhancer's orientation relative to the promoter. Moreover, repression levels observed for weakly repressed synthetic enhancer circuits reflect the number of TetR-binding sites on the cassette by yielding discretely separated values for each cassette type. This is highlighted by the colored lines, which denote each of colored data sets representing the repression functions $r_1(L)$, $r_2(L)$, and $r_3(L)$ (see "Theory: Model for Looping Initiated Transcription" in the Supplemental Information for the functional form of these terms) on the graph.

To understand the origins of regulation at a distance in our synthetic enhancer system, the thermodynamic model framework tells us how to go beyond the two-state description introduced in Figure 1. In particular, we have to account for all of the different states of occupancy in which TetR can be bound to the DNA looping region. To that end, we add an additional set of states to our thermodynamic modeling framework, which provides a convenient scheme for characterizing the different states of the promoter and their relative probabilities. As shown in Figure 2C for the 1-Tet case, the model now has four states that come in two broad categories: unlooped and inactive and looped and active, each with and without TetR bound. Unfortunately, our knowledge of the geometric details of the loops in the repressed case (i.e., when the cassette is bound by TetR proteins) is too meager to adopt a "first principles" approach, which would allow us to relate the looping capacity in the presence of TetR to the looping capacity in its absence. As a result, the states and weights are still written in terms of the looping capacity, but now the looping capacities themselves are undetermined parameters.

However, for the long looping length limit ($L >> L_t$), simple polymer models can be used to develop intuition for the resulting repression (Phillips et al., 2009). Using these theoretical results and the model presented in "Theory: Model for Looping Initiated Transcription" in the Supplemental Information, we can derive an expression for long-distance repression that is a ratio of the repressed to the unrepressed looping capacity functions (Equations S23 and S24), which converges to a fixed value and gives a sense of the theoretical underpinnings for r(L). Consequently, at the very long loop length limit, both the model and experiment indicate that these repression values (denoted by the dashed lines) seem to converge on a particular constant for each cassette configuration, rather than approach the nonrepressed value of 100%.

Using the long-looping length limit and the repression values observed for the strong repression regime, we can approximate the functional form of the repressed looping capacity functions (fits in Figure 2B) for each cassette using the same functional form exploited earlier. Using these functions, the data can be compactly represented by a simple function that is consistent with both the transition lengths (L_t) and the saturation values that appear to be correlated with the number of TetR-binding sites and the distance between the beginning of the NRI#1 site and the last TetR-binding sites (see Figure S4).

Multiple TetR-Binding Sites Generate Step Functions from a Variable Input

The long-range repression capability of our synthetic enhancer system discussed above has further regulatory potential. This observation suggests a design strategy for constructing synthetic enhancers. By tuning the concentrations of an input signal, which alters the binding probability of the regulatory proteins, the level of gene expression can, in turn, be systematically tuned between different discrete values. In the case of TetR, this can be done simply by titrating variable amounts of a soluble ligand anhydrous-Tetracycline (aTc), which prevents the binding of TetR to its binding site by inducing a conformational change (Orth et al., 2000).

We studied the regulatory output of four different types of binding site cassettes—1-, 2-, 3-, and 6-Tet—in response to the variable input signal. In order to compare the output functions for the different cassettes, we plot the data (Figure 3) by constructing a ratio of the fluorescence level measured in the presence of a given ligand concentration divided by the maximal average fluorescence level (i.e., when the cassette is most likely unoccupied by TetR at saturating concentrations of aTc; labeled 100% on the plots).

In Figure 3A, the regulatory function for the 2-tet cassette is presented. We observe a response that exhibits three discrete values of expression: a repressed state, a sharp transition at ≅10 ng/ml aTc to an intermediate partially repressed level, and a final transition at ≅200 ng/ml aTc to an unrepressed expression level.

In order to understand the intermediate expression level of the regulatory output function, we constructed two additional synthetic enhancers. These enhancers were constructed with identical looping lengths to the 2-Tet enhancer and contain only a single binding site for TetR at either the distal or proximal binding site location of the 2-Tet construct. Examination of Figure 3B shows that the weak repression level ($r_1(L)$) measured for the single binding site cassettes is in reasonable accord with the intermediate level of the repression ratio in Figure 3A and with the weak repression regime for the 1-Tet cassette repression data (Figure 2B). Therefore, it is likely that the intermediate level observed for the 2-Tet enhancer reflects the partial TetR occupancy configuration (Figure S4A) for the two-binding site architecture.

The regulatory output function for the 3-Tet cassette shown in Figure 3C also exhibits a series of discrete expression levels. In particular, this case is characterized by four values: a fully repressed state and a sharp transition at 10 ng/ml to a set of three nearby expression levels that are located at values of roughly 70%–80%, 90%, and 100%, respectively. Alternatively, one may choose a more conservative interpretation of the data shown in Figure 3C as having a single intermediate level at ≅70%–80% and a shallow increase to 100% repression ratio thereafter.

The 3-Tet output function can be understood qualitatively using similar logic to that introduced in thinking about the 2-Tet cassette regulatory function. For this case (Figure 3D), there is one configuration for full occupancy, one for an unoccupied state, and three configurations each for single and double occupancies. To show that the steps shown in Figure 3C reflect these partial occupancy states, we measured the repression values for six additional cassettes that account for all possible occupancy configurations (Figure 3D). We found that only the triply occupied configuration is strongly repressed, whereas the other configurations are weakly repressed with values of 40%–45% and 60%–80% of full expression for double and single occupancy, respectively, thereby supporting the idea that the discrete jumps in the repression ratio levels are associated with either the single or double occupancy configurations. Interestingly, the repression ratio value of the first (and perhaps only) intermediate coincides approximately with the average repression level ($r_1(L)$; purple shade; Figure 3D) of the three single occupancy configurations. This indicates that the dominant state at these aTc concentrations is the single occupancy configuration.

The next step in the progression of increasingly complex enhancer architectures corresponds to a case with six TetR-binding sites. The regulatory output function (Figure 3E) does not exhibit an increase in the number of intermediates but instead is characterized by two intermediates with more evenly spaced repression ratio values and with sharper transitions that produce a more distinct step-like function than for the 2- and 3-Tet cassettes (see also Figure 4). Here, the first intermediate repression ratio state is located at 65% of the unoccupied cassette maximum and the second at 75%–80% of the maximum. These values are markedly different from the 80% and 90% values that were measured for the 3-tet cassette.

Combinatorial Control in a Synthetic Enhancer

Examining the data for the 1-, 2-, 3-, and 6-binding site cassettes more closely, we find additional regulatory features that likely would not have been guessed a priori. The dose-response for each TetR cassette type indicates that the transition

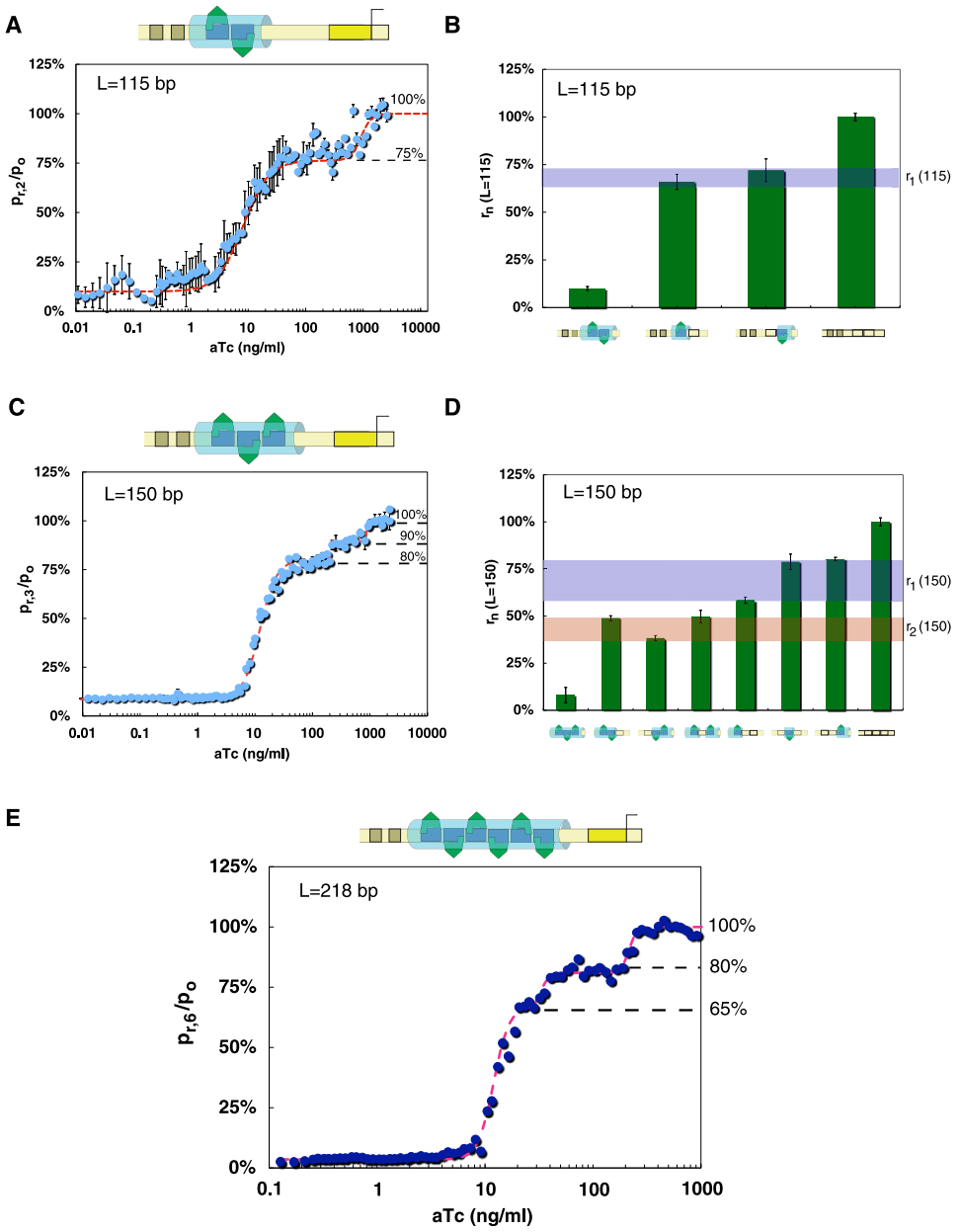

Figure 3. Synthetic Enhancers Convert Variable Ligand Input to Discrete Output Step Function

(A) High-resolution titration in 48-well plates of aTc with a 2-tet cassette at L = 115 bp. The data show three discrete states separated by transitions.

(B) Repression levels measured for synthetic enhancers characterized by a deletion of either one or both of the TetR-binding sites at L = 115 bp. The purple shading corresponds to the weak repression value r_1 (L = 115 bp).

(C) 3-Tet repression ratio at L = 150 bp exhibiting four discrete states, with the upper three closely clustered at average repression ratio values of ~80%, ~90%, and ~100%.

(D) Repression levels measured for synthetic enhancer cassettes (L = 150 bp) containing zero, one, or two TetR-binding sites arranged in configurations that mimic the three binding site enhancers' partial occupancy states due to aTc titrations. The purple and orange shading corresponds to the weak repression values r_1 (L = 150 bp) and r_2 (L = 150 bp).

(E) Data for the 6-Tet cassette showing only four states, characterized by increased separation and sharper transitions between the intermediate states.

The dashed red lines in (A), (C), and (E) correspond to empirical fits of two (A) or three (C and E) Hill functions stitched together in a piece-wise continuous fashion that highlight the transitions and levels observed in the data. Error bars correspond to the standard deviation from multiple measurements. See also Figure S5.

(Figures 4A–4D) between the low repressed state and the first intermediate are characterized by an increasingly steeper transition that can be empirically quantified by a Hill coefficient greater than one. Interestingly, the Hill coefficients that were extracted turn out to be roughly equal to the number of TetR-binding sites. This result seems to imply that the regulatory function reflects an

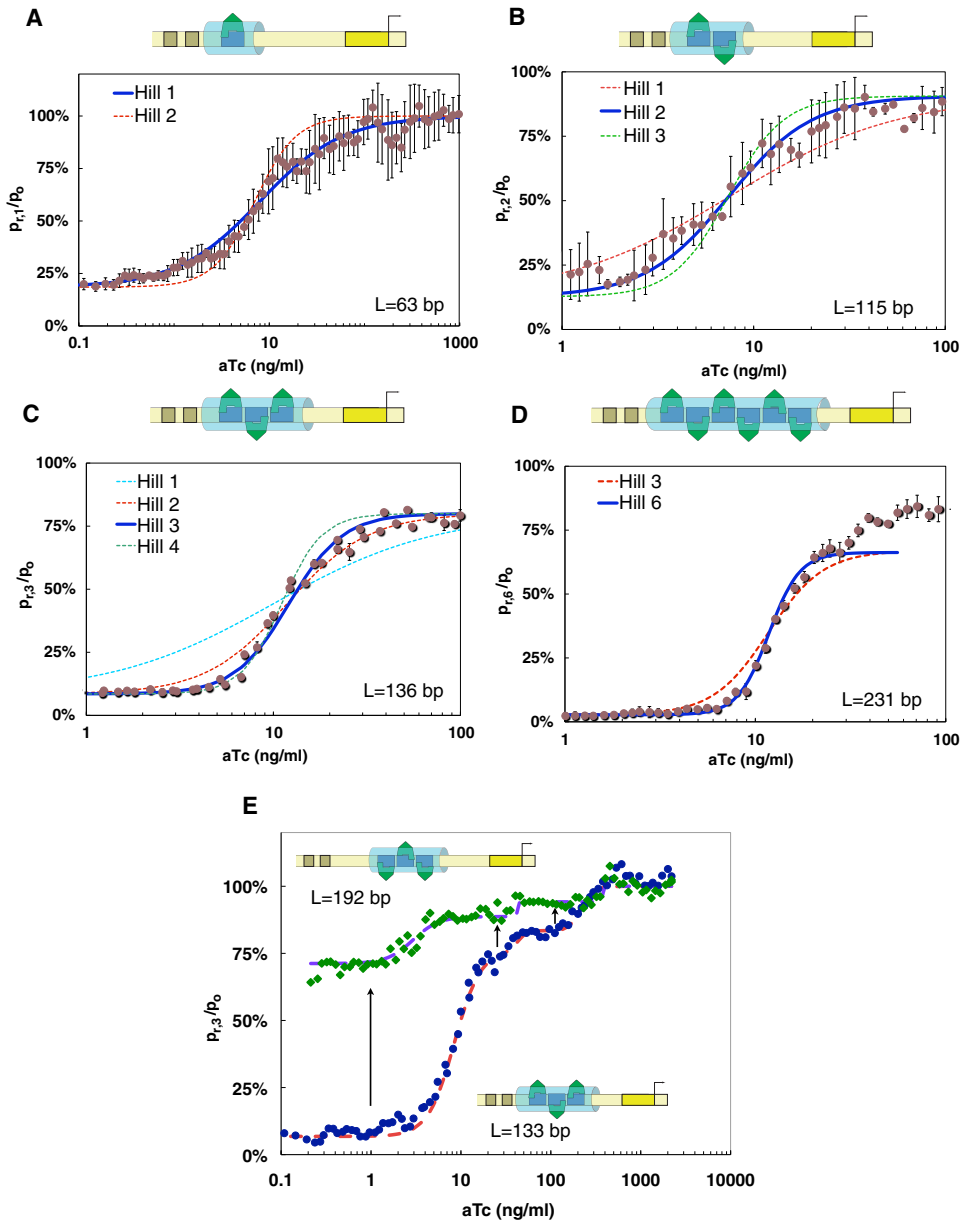

Figure 4. Coding and Computational Characteristics of Synthetic Enhancers

(A–D) Transition from the strongly repressed state to first intermediate level. The transition in all Tet cassettes (1, 2, 3, and 6) is best fitted by a Hill function of order (n), which roughly equals the number of binding sites. Dashed lines in each curve signify fits with Hill functions of n + 1 or n − 1, typically showing that only Hill functions of order n fit the data well.

(E) By shifting the cassette toward the σ^{54} promoter and away from the driver NRI#1 and #2 sites, a similarly shaped regulatory function (top) is observed. Error bars correspond to the standard deviation from multiple measurements. See also Figure S5.

effective interaction in the factors that bind to the cassette, which can be interpreted as a form of molecular counting.

To further examine the mechanistic underpinnings of our measurements, we examined the output function of additional synthetic enhancers with the binding site cassettes moved upstream a larger distance from the end of the NRI#2 site. This serves to further explore the effects of looping modification on the regulatory output and also as a control for whether or not our placement of the binding site cassette 28 bp upstream of the NRI#1,2 sites interferes in some nontrivial fashion with the binding of NRI∼P. Figure 4E shows that, for a synthetic enhancer with the three TetR-binding site cassette placed 45 bp downstream of the end of the NRI#2 site, the output function keeps its elementary characteristics (i.e., a strongly repressed state, a transition to one or two weakly repressed or unrepressed states, and transition steepness characterized by

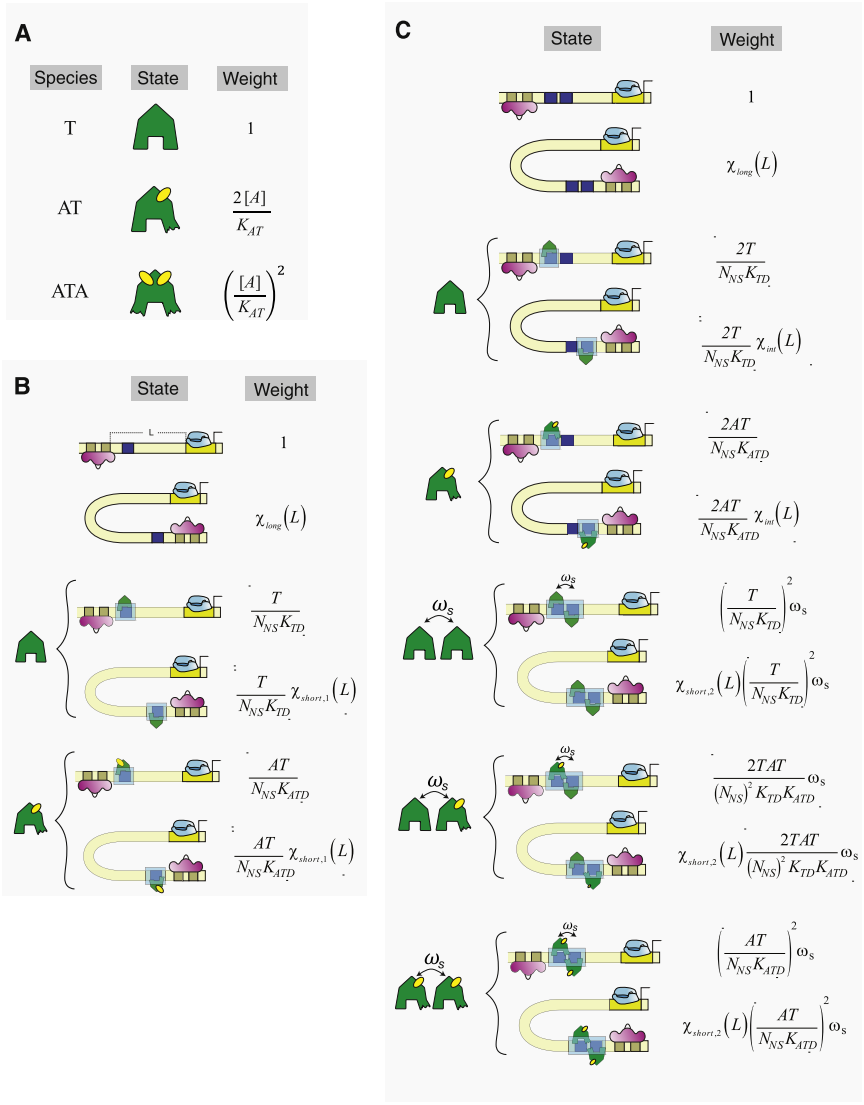

Figure 5. Generalized Model Schematic for Repression Ratio Data

(A–C) The models and their corresponding states and statistical weights are shown for (A) the interaction between aTc and TetR in solution, (B) the states and weights used for computing the repression ratio model function for the cases of a single TetR, and (C) two TetR. In (B) and (C), we now include states with the single aTc-bound TetR form. This protein has a binding affinity to the specific binding sites of TetR, which is two to three orders of magnitude lower than the free form of TetR. Furthermore, the two TetR model in (C) has a new parameter ω_s, which describes the interaction between adjacent TetR molecules. This interaction is crucial for the formation of steps in our model.

See also Figure S5 for model fits.

a Hill coefficient of three) regardless of where the cassette is positioned within the spacer region. Thus, the results shown in Figure 4E and the different response functions for the 1-, 2-, and 6-Tet cassettes (see Figure 4A, Figure 3A, and Figure 3E, respectively) suggest that each cassette type apparently encodes a particular output function, whose characteristic dose-response output depends on the geometry and binding site arrangement of the various TetR-binding cassettes and a possible interaction between TetR proteins bound on the cassette.

Modeling the Enhancer Output Functions

Given the modeling framework discussed in "Theory: Model for Looping Initiated Transcription" in the Supplemental Information, which were used to model the looping and the bimodal repression data, is it possible to generalize this scheme to reproduce the output functions shown in Figure 3 and Figure 4? In order to address this question, we need to develop a proper mechanism by which to extend the thermodynamic model to account for the aTc titrations. In doing so, we incorporate the following assumptions: the observation (Lederer et al., 1995) that up to two aTc ligand molecules can bind a single TetR dimer and that TetR can bind its DNA-binding site in two forms: unoccupied and occupied by a single aTc ligand but with different K_ds (see Table S4). These assumptions are based on crystal structure analysis (Orth et al., 2000) and in vitro binding experiments (Lederer et al., 1995, 1996). In the former, the ligand is shown to increase the distance between the DNA-binding motifs on the dimer, thus reducing the binding affinity to DNA of a protein bound by a single ligand and abolishing it altogether when both ligands are bound. In the latter, binding curve analysis suggests that more than one bound ligand is required to abolish TetR binding to the DNA.

These assumptions allow us to formulate states and weights prescriptions (see schematic in Figure 5A), which generate mathematical expressions (see "Theory: Model for Enhancer Repression via Induction" and Equations S32–S34 in the Supplemental Information) for the number of TetR molecules in various states of aTc occupancy—*T*, *AT*, and *ATA* corresponding to the number of free TetR proteins, TetR occupied by a single molecule of aTc, and doubly occupied TetR, respectively. Given this relationship between TetR and aTc, we were then able to install those results into our states and weights schemes for the various enhancer occupancies, which in turn allowed us to formulate a model for the repression ratio data (Figures 5B and 5C for generalized model schematics), which not only accounts for the looping size effect due to TetR binding, but also illustrates how this binding is altered in the presence of different concentrations of aTc.

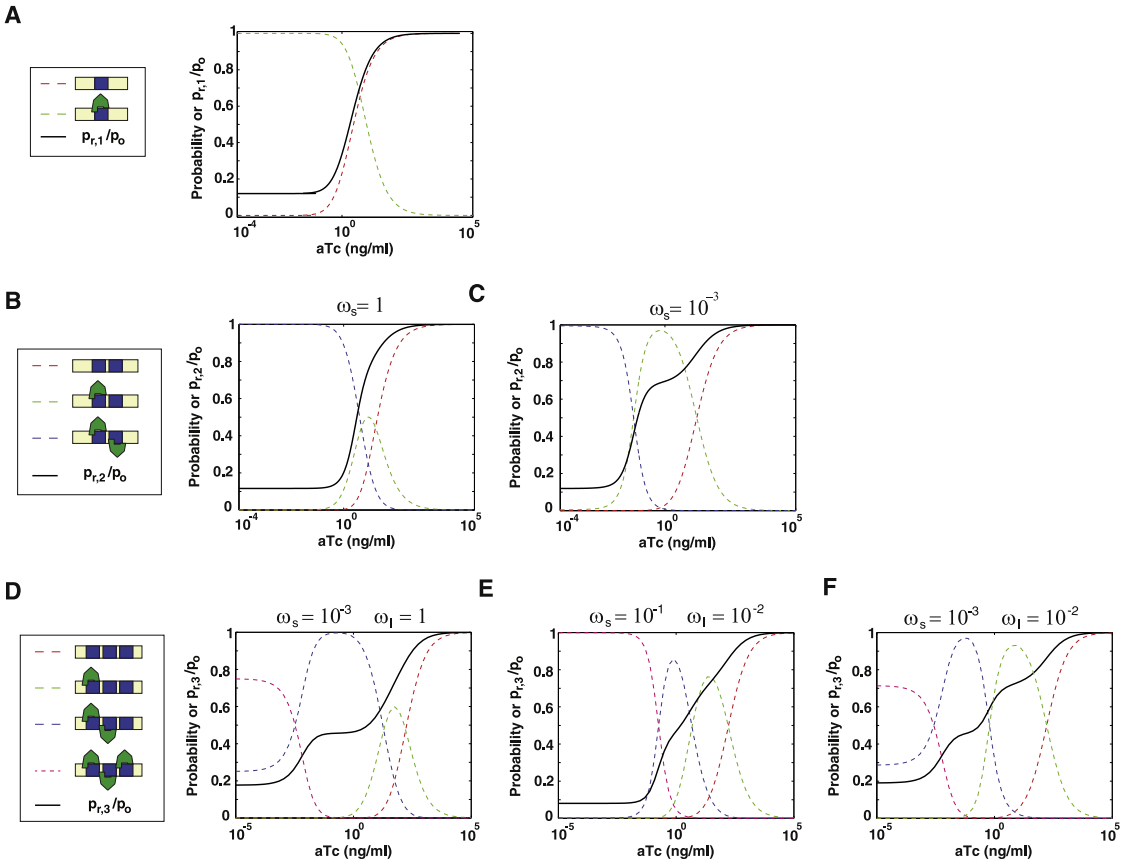

Figure 6. Theoretical Repression Ratio Curves and Associated Probabilities

In all panels the red, green, blue, and violet dashed lines correspond to the no occupancy, single, double, and triple occupancy state probability distributions respectively plotted as a function of aTc concentration. The thick black line corresponds to the theoretical repression ratio dose-response function computed at each aTc concentration from the individual probability distributions.

(A) Occupancy states and dose-response function for the single binding site case.

(B) Occupancy probability states of the two TetR-binding site model and associated dose-response function constructed using the parameters from (A) and $\omega_s = 1$.

(C) Same model as described in (B) but with $\omega_s = 10^{-3}$, implying that it is energetically unfavorable to have two TetR molecules bound next to each other.

(D–F) Occupancy probability states and associated dose-response functions for the three TetR-binding site model for cases in which the short- and long-range interaction parameters take the values (D) $\omega_s = 10^{-3}$ and $\omega_l = 1$, (E) $\omega_s = 0.1$ and $\omega_l = 10^{-2}$, and (F) $\omega_s = 10^{-3}$ and $\omega_l = 10^{-2}$.

Sample fits of the model to the data sets in Figure 3 and Figure 4 are shown in Figure S5.

First, we consider a model for the 1-Tet cassette. Figure 6A shows a typical repression ratio curve and associated occupancy state probabilities that can be obtained for a wide array of parameter combinations. The model for the 1-Tet case captures the essential features of the 1-Tet data (Figure S5A), as does the empirical fit given by a Hill function with Hill coefficient one, as shown in Figure 4A.

In order to extend the model to the 2-Tet case (see "Theory: Model for Enhancer Repression via Induction" in the Supplemental Information), we incorporate an additional parameter (ω_s) that accounts for any interaction that may be incurred between bound proteins on neighboring TetR sites. If this parameter is less than one, then the bound proteins exhibit anticooperative behavior, which leads to increased stability for the single occupancy configurations as compared with the double occupancy one. On the other hand, if ($\omega_s > 1$), then this parameters amounts to a cooperative interaction, which leads to a preference for the doubly occupied state as compared with other cassette occupancy states (data not shown).

In Figures 6B and 6C, we plot the individual probabilities (Equation S47) for the cassette suboccupancies as a function of ligand concentration for the 2-Tet case for two values of (ω_s): 1 and 0.001. The blue dashed lines in both panels correspond to the double occupancy probability, which approaches one for very low ligand concentrations and declines sharply thereafter. Likewise, the red lines correspond to the no occupancy configuration, and as expected, the probability of this state approaches one for very high ligand concentrations. The single occupancy probability (green lines) varies sharply between both panels. For values of ($\omega_s \cong 1$) (Figure 6B), it overlaps significantly with the other two probabilities, leading to a relatively small overall contribution from the single occupancy

configurations, which results in an output function that lacks an intermediate step (Figure 6B, black line). However, for values of (ω_s) that promote anticooperativity in the protein-protein interaction, the overlap of the probabilities is significantly reduced (Figure 6C), which in turn leads to an intermediate step in the output function. Thus, according to our model, the reduced stability of the double occupancy configuration is critical for the formation of the step function.

Extending the model further to the 3-Tet case (Figures 6D–6F and Figure S5C for fits) and varying the value for (ω_s) leads to the emergence of a step function for decreasing values of ω_s characterized by a single intermediate, as for the 2-Tet case. The plot in Figure 6D shows a clear signature for a step at a repression ratio level of $\cong$0.4–0.5, with a second additional sharp transition to the top level corresponding to the unoccupied cassette configuration. For slightly lower values of ω_s, the model produces an output function (Figure 6E) that looks similar to the data in Figure 3C. However, no matter what value of ω_s is chosen, the model is unable to produce two intermediate states. In order to generate a step function with two intermediates (Figure 6F), one has to introduce a second weaker anticooperativity term (ω_l) for the next to nearest neighbor interaction. As a result, we conclude that the existence of anticooperativity interaction parameters seems to be a crucial feature of any model that attempts to reproduce the particular discrete output functions obtained by the experiments, with the number of intermediates steps reflecting the extent of the protein-protein interactions (i.e., nearest neighbor, next-nearest neighbor, etc). However, a full microscopic understanding of the function of these synthetic enhancers requires a deeper knowledge of both the DNA mechanics and the ways in which the repressors interact both with each other and with their DNA substrate.

Conversion of the σ^{70} Activator TraR to a Repressor Using Synthetic Enhancers

We reasoned that there was nothing special about the character of TetR as a DNA-binding protein that led to the observed behavior of our synthetic enhancer. To the extent that this hypothesis is correct, we should be able to replace TetR with some other DNA-binding protein and obtain a qualitatively similar regulatory output. To that end, we constructed additional synthetic enhancer cassettes containing binding sites for the activator TraR. In particular, under normal circumstances, TraR, a LuxR homolog found in *Agrobacterium tumefaciens*, acts as a transcriptional activator of σ^{70} promoters. In *E. coli*, however, its transcriptional activation capability is abolished, though the specific DNA-binding activity remains (Qin et al., 2009 and references within). Thus, in our case, we can use this protein in the enhancer context to alter the looping region just as we did with TetR.

The results obtained previously for the TetR systems (Figure 2, Figure 3, Figure 4, Figure 5, and Figure 6) indicate that the behavior of the output functions that are generated by the class of models presented here depends strongly on three parameters: the values of the looping capacities for the different enhancer states of occupancy by the enhancer binding protein (Figure 7A), the number of binding sites (Figure 7B), and the protein-protein interaction parameter (Figure 7C). In particular, the protein-protein interaction parameter determines whether the regulatory output will exhibit a smoothly decreasing expression level function ($\omega_s \cong 1$) or be characterized by sharp transitions and an intermediate expression level step ($\omega_s << 1$). Because the presence of a step in the regulatory output function indicates that the states with several enhancer-binding proteins bound are relatively unstable, the model predicts that this effect is attainable experimentally if a large mutual exclusion effect is engineered into the synthetic enhancer design.

Due to the fact that the DNA binding probability for TraR increases as a function of ligand concentration (see Figure 7D and "Theory: Model for Enhancer Repression via Induction" in the Supplemental Information), the model predicts that it is possible to obtain a regulatory output function that is qualitatively a mirror image of the output function obtained for the synthetic enhancer architecture with three TetR-binding sites (for states and weights, see Figure S6). Consequently, we opted to design the TraR synthetic enhancer with 6 bp spacing between the binding sites to ensure that a mutual exclusion effect will be present as a result of presumed excluded volume effects between the bound TraR dimers. Figures 7E and 7F show the experimental results and model predictions. At low ligand concentrations of the small inducer molecule that is necessary for TraR to bind to DNA, N-(3-oxo-octanoyl)-L-homoserine (3OC8), the enhancer regulatory response is characterized by a small magnification ($\cong$7%) of expression levels as compared with the unoccupied enhancer for 3OC8 concentrations that are less than 10 nM. For larger concentrations, repression characterized by clearly detectible steps is observed with a minimal value of $\cong$60% of the unoccupied enhancer expression level. The data indicate that a well-separated intermediate in repression values occurs at $\cong$90% of unoccupied expression level and ranges from $\cong$30 to 500 nM in 3OC8 concentration, validating the model's qualitative predictions and our general approach for inducing regulatory response in synthetic enhancer design.

DISCUSSION

We explored transcriptional and regulatory characteristics of an enhancer-based transcriptional system by constructing increasingly complex enhancer elements from the ground up. Our approach was predicated on the assumption that a bacterial enhancer can be constructed as a modular object made of three connected components: driver-binding sites, a poised σ^{54} promoter, and small DNA cassettes containing several binding sites for DNA-binding proteins. In this work, we restricted ourselves to using the same module for the driver and poised promoter while varying the enhancer-binding protein binding site module. However, we suspect that any of the other modules can be altered to access an even richer space of regulatory effects.

We then proceeded to characterize our synthetic enhancers' regulatory output functions using experimental measurements and a set of thermodynamic models. Our results show that, unlike the conventional model for repression, wherein a repressor inhibits transcription by competing

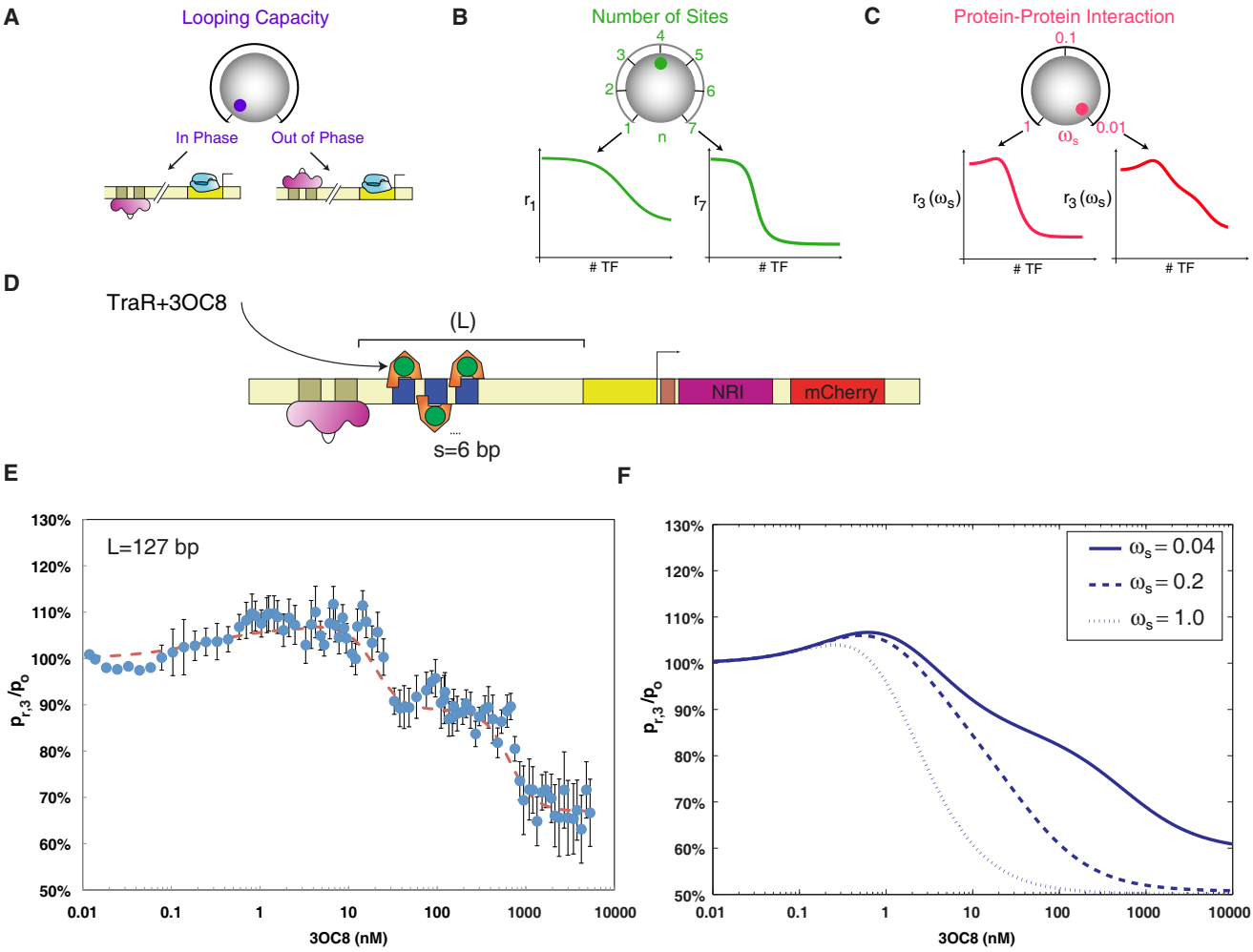

Figure 7. Prediction and Regulatory Behavior of an Enhancer Designed from the Ground Up

(A–C) Schematic representation of the different experimental knobs for controlling enhancer regulatory output.

(D) Circuit schematic for the TraR synthetic enhancers with three TraR-binding sites arranged with 6 bp between each site (as compared with 16 bp used for TetR). The cartoon for the TraR protein signifies that only the dimeric isoform of TraR bound to the cognate ligand 3OC8 can bind DNA.

(E) Regulatory dose-response function for the TraR synthetic enhancer over six decades of 3OC8 concentration. The dashed red line corresponds to an empirical fit of two Hill functions stitched together in a piece-wise continuous fashion.

(F) Model prediction (Equation S50) for regulatory output of the TraR synthetic enhancer showing examples with three values of the short-range interaction parameter. We used the following normalized looping capacity values (i.e., each value is divided by χ_o) for all three curves: $[\chi_L, \chi_{int1}, \chi_{int2}, \chi_{short,3}] = [1, 1.15, 0.85, 0.5]$. See Figure S6 for a detailed graphical representation of the statistical states and weights for this model. See also Figure S7 for schematic of naturally occurring bacterial enhancers.

for the RNAP-binding site or by interfering with RNAP initiation, the synthetic enhancers exhibit repression by a modification of the DNA's capacity to loop. This leads to a regulatory output that is characterized by two key modes: a strongly repressed state in which the enhancer is unlikely to loop and a weakly repressed state in which looping is more likely at short and long looping lengths, respectively. Within each mode, the resultant level of repression depends on the enhancer element properties (i.e., number of binding sites, transcription factor binding regions, binding site arrangement and spacing, etc.) and weakly on the length of the loop (Figure 2B). Therefore, these results provide a mechanistic model for regulatory action at a distance by showing that regulatory effects can be systematically generated when the transcription factors are bound at large distances (i.e., hundreds of bps) from the basal promoter.

One striking outcome induced by the various repression states observed for our synthetic enhancers is the emergence of step-like dose-response regulatory output functions. In the Results section, we showed that the steps that form in the response for the 2-Tet, 3-Tet, and 6-Tet cases can be explained by repression levels of preferred cassette occupancy states. The preferred states, in turn, are determined by various anticooperativity parameters, which are used to model a destabilizing interaction between two TetR proteins that are bound in the vicinity of one another.

Given this analysis, we then asked whether it is possible to utilize these underlying mechanisms that are responsible for enhancer regulatory output and design a new synthetic enhancer from the ground up with a predetermined output function using a completely different enhancer-binding protein. We showed that, if we replace the TetR protein by another DNA-binding protein (TraR) and conserve the binding geometry (i.e., proteins are bound in opposite orientation with spacing of 6 bp for TraR and 16 bp for TetR), the same step-like regulatory response is observed in accordance with the model's qualitative predictions. As a result, our data suggest that the specific identity of the enhancer-binding protein (TraR and TetR are generic choices of DNA-binding proteins) is not as crucial to the regulatory output as the arrangement and number of its binding sites. Consequently, the design of enhancer regulatory output is reduced to a consideration of the variable looping geometry induced by the presence of DNA-binding proteins within the loop.

The observed discrete levels of the regulatory output (Figure 2) and the transitions between steps (Figure 3,4,7) of this output illustrate that a form of molecular counting is taking place at the synthetic enhancer. Because regulation has traditionally been used to explain the phenomenon of gene switching from "on" to "off" and vice versa, how do we then classify cases like that described here, wherein there are apparently more than two discrete regulatory states that can be accessed within a singular regulatory motif?

The regulatory effects observed with our synthetic enhancers can be interpreted via our model as a cumulative outcome of three analog knobs individually tuned to particular values (Figures 7A–7C). These knobs are the looping capacity values, the number and arrangement of transcription factor binding sites, and the character of the protein-protein interaction. All three of these tuning variables are distinct yet affected by the particular state of the others. For instance, we showed that the ability to loop is affected by the presence or absence of DNA-binding proteins and by the number of binding sites. Furthermore, the number of bound proteins for a given concentration of inducer is, in turn, affected by the protein-protein interaction parameter, which reflects the number of active proteins present in the cell.

Even though our experiment and model allowed us to conveniently identify or isolate these control parameters, at present, the models serve primarily as a conceptual framework for understanding the behavior of the synthetic enhancers as a function of the various regulatory knobs that can be tuned. Unfortunately, for the time being, it is not possible to predict either the looping capacity or the protein-protein interaction parameters from first principles. In particular, for the cases presented here, we showed that the looping capacity can be repressive for the case of TetR or repressive and activating for TraR. Both of these observables are apparently related to the particular localized protein-DNA interactions, yet we are unable to formulate a first principles theoretical model for these quantities. These uncertainties are an inheritance of our current limited understanding of in vivo DNA mechanics, protein-DNA interactions, and protein-protein interactions for neighboring transcription factors. At the same time, we view the kind of interplay between experiment and theory played out here as precisely the type of approach that will allow us to begin to develop quantitative intuition for all of these phenomena.

Given these limitations, what are the practical lessons learned from our synthetic enhancer's capability to count molecules or "measure" cellular concentration of proteins? Recently, molecular counting was demonstrated using gene regulatory networks via both systems (Long et al., 2009) and synthetic biology (Friedland et al., 2009) approaches. When comparing these two examples, we find that they describe two different forms of counting. In Friedland et al. (2009), the authors demonstrate a chemical pulse counter, which yields a singular output once a particular pulse number is reached. On the other hand, the quorum-sensing counter shown by Long et al. (2009) generates an output expression level, which is a discrete function of the number of inputs integrated (in their case, two). The behavior of the circuits that we have constructed are analogous to integrative counters but exhibit a capability to integrate more than two inputs in a compact DNA sequence architecture. As a result, it is tempting to speculate that gene regulatory circuits, which utilize enhancers as input integrators, can therefore enable an enriched regulatory potential.

Finally, the motivation for building synthetic enhancers from the ground up is to not only generate some complex regulatory phenomenon, which in this case tests our understanding of protein-DNA interactions and poised transcription, but to also try to isolate underlying mechanisms that are responsible for natural regulatory phenomenon. Similar constructionist approaches have been used often in recent years to study gene regulatory networks, and in the many examples published to date (e.g., Basu et al., 2005; Elowitz and Leibler, 2000; Gardner et al., 2000), gene circuits synthesized de novo often yielded important insights into the underlying mechanism of protein networks in biology. Hence, the question remains of whether any of the above results and their interpretations provide new insight into regulatory phenomena observed in natural bacterial enhancers.

As an example of natural bacterial enhancers, the wild-type NRI∼P system in *E. coli* contains three additional NRI sites (#3–#5) (see Figure S7 and "Theory: Model for Enhancer Repression via Induction" in the Supplemental Information) that flank the #1 and #2 sites and σ^{54} promoter, in what we defined as the looping region (see Figure 1). Deletion of these sites (Atkinson et al., 2002) has been shown to increase expression in discrete amounts driven by the hexamer bound at the #1 and #2 sites. These additional sites have been dubbed "governor sites" as a tribute to the fact that they limit or inhibit the overall expression level. Thus, we can effectively consider this natural system as analogous to the synthetic enhancer considered here with a "cassette" of three additional NRI-binding sites.

To explore this analogy further, we examined the binding site architecture of three additional bacterial enhancers (http://regulondb.ccg.unam.mx/) (Figure S7). In a manner similar to the synthetic enhancers, these natural enhancers form entities that are capable of integrating multiple inputs upstream of a poised σ^{54} promoter. The binding site architectures imply that the regulatory output exhibited by these enhancers may be characterized by a similar modeling approach to the one used here. Because the ingredients used to construct our synthetic

enhancer are all common architectural elements in real transcriptional networks, we argue that the capacity to assemble these elements as done here can provide a predictive model for deciphering the regulatory output of additional bacterial enhancers in the natural context as well. Given these similarities, it is tempting to speculate that the modification of the looping capacity mechanism explored in our work might actually be a strategy adopted for the regulation of natural enhancers in bacteria.

EXPERIMENTAL PROCEDURES

Synthetic Enhancer Cassette Design

Synthetic enhancer cassettes (Table S1) were designed as follows. First we computationally designed 100 bp sequences that had a minimal probability to bind DNA-binding proteins. This was done by constructing an algorithm that randomly generated a set of 1 million 34 bp sequences. The sequences were compared to the roughly 1900 known specific DNA-binding sites for *E. coli* transcription factors obtained from RegulonDB (http://regulondb.ccg.unam.mx/). Each calculated sequence was scored by first computing the percent homology with a particular binding site, weighting that number by an exponential weight that heavily favors low homologies, and finally totaling the values obtained for each of the 1900 binding sites (sequences that matched a known binding site were eliminated). After obtaining the sequences with the lowest scores, a second run was carried out on the complementary sequence of the lowest-scoring 1% of the original sequences. The scores of the two runs were combined, and sequences with the lowest combined scores were listed in order. The sequences were predominantly GC rich (~75%) with very low A and T content (~25%). We ordered the spacer92 (Table S1) sequence using two complementary primers (IDT).

Cassettes containing TetR-binding sites were designed as follows (all containing a tandem of NheI sites). The 1-Tet cassette included the high-affinity (10 pM) TetO2 site (Hillen and Berens, 1994) (Table S1). The 2-Tet cassette included the TetO1 site (30–50 pM) site (Hillen and Berens, 1994), a 16 bp spacer (obtained from the calculated spacer sequence; see above), and a TetO2 site. The 3-Tet cassette contained two TetO2 sites, two spacer sequences of 16 bps (determined using the above algorithm), and one TetO1 site. The 3-Tet-S cassette has additional spacer sequences placed in front of the first TetO1 site and after the last site. The 6-Tet cassette is effectively a double cassette made of a tandem of 3-Tet cassettes.

All cassettes were ordered as complimentary oligos from IDT. Oligos were hybridized as follows (in saline solution containing 10 mM $MgCl_2$) and then placed on ice: 2′ @95°C, 15′ @65°C, 5′ @42°C. Hybridized dsDNA cassettes were gel purified and digested with NheI before being used as an insert in the cloning step.

Looping Length Dependence Assay

20 ml of fresh LB with appropriate antibiotics was inoculated in 125 ml flasks with overnight starters of synthetic enhancer strains characterized by different looping lengths (i.e., 3.300LG cells + synthetic enhancer plasmid + p3Y15 plasmid; Atkinson et al., 2003). Cultures were vigorously shaken at 37°C (Innova), and fluorescence measurements were taken at 30 min intervals for roughly 5 hr to cover the midlog growth range. For each measurement, 200 μl of culture was dispensed in each of four wells of a 96-well plate (Corning Costar–Fisher Scientific). The 96-well plates were read by a plate reader (Tecan–Infinite 200) at 580/610 excitation/emission with gain 100 and appropriate controls for autofluorescence and glnAp1 leakage. The fluorescence results for the four wells were averaged and normalized by a reading of the culture's OD600. S/N was > 10 for all synthetic enhancer strains tested with respect to leakage and > 20 with respect to auto fluorescence (obtained from a null strain).

Repression Measurement Assay

Repression level measurements were carried out as follows: first, synthetic enhancer plasmids were transformed with either pACT-Tet (Figure S1A) or pACT-Tra plasmids in 3.300LG (Atkinson et al., 2003) cells (in which the *traR* gene replaces the *tetR* gene). Next, synthetic enhancer strains were grown in fresh LB with appropriate antibiotics (Kan/Amp) to midlog range, as measured by a spectrophotometer (Pharmacia Biotech) OD600 of ≅0.6 and were resuspended in low-growth/low-autofluorescence BA buffer (for 1 l – 0.5 g Tryptone [Bacto], 0.3 ml Glycerol, 5.8 g NaCl, 50 ml 1M $MgSo_4$, 1 ml – 10 × PBS buffer – pH 7.4, 950 ml DDW). 1 mM IPTG was added to induce the circuit at this point to deactivate the LacI protein that represses the glnAp2 promoter. 2 ml of resuspended culture with IPTG were dispensed in each well of a 48-well plate. The plates were then incubated in a 37°C shaker until cultures reached growth steady state. Measurements of fluorescence levels were taken by dispensing 200 μl of culture in each well into a 96-well plate and were carried out on a plate reader as mentioned above. All repression measurements were done in triplicates with cultures grown from individual synthetic enhancer strain colonies.

To get the percentage of inhibition, autofluorescence levels were subtracted from expression levels measured for strains with and without endogenous TetR. Subsequently, the ratio of the adjusted fluorescence level for the +TetR strains to the –TetR strains was taken.

Repression Ratio Measurement Assay

Synthetic enhancer strains containing the pACT-Tet or pACT-Tra plasmid were initially grown in LB, resuspended in the low growth buffer, and dispensed in the 48-well plates. In this case, appropriate concentrations of aTc or 3OC8 (sigma) were dispensed in each well, spanning four to six orders of magnitude. For each strain, we used two plates to allow for 94 different readings of fluorescence as a function of aTc concentration (two wells were used as –IPTG controls). We carried out each measurement in duplicates, i.e., four plates per measurement.

To compute the repression ratio levels as a function of aTc or 3OC8 concentrations, each fluorescence ratio value was calculated using a running average algorithm. This entails averaging three to five raw fluorescence readings for every fluorescence value shown, whereby the averaging is carried over adjacent inducer concentrations. This algorithm is used to smooth out short-range fluctuations and highlights the large-scale features that span wide concentration ranges.

Strain Construction

See Extended Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and five tables and can be found with this article online at doi:10.1016/j.cell.2011.06.024.

ACKNOWLEDGMENTS

We would especially like to thank Prof. Frances H. Arnold for providing lab space and the forum to conduct thorough discussions as this project was evolving. We would also like to thank Eric H. Davidson for important early discussions and Alex J. Ninfa for plasmids and strains. R.A. was supported by a NIH Ruth L. Kirschstein fellowship, a Caltech CBCD grant, and the NIH through award NIH ARRA R01 GM085286-01S. R.P. and H.G.G. gratefully acknowledge awards NIH ARRA R01 GM085286-01S, R01 GM085286, and the NIH Director's PIONEER Award DP1 OD000217.

Received: August 24, 2010
Revised: January 25, 2011
Accepted: June 14, 2011
Published: July 7, 2011

REFERENCES

Atkinson, M.R., Pattaramanon, N., and Ninfa, A.J. (2002). Governor of the glnAp2 promoter of Escherichia coli. Mol. Microbiol. *46*, 1247–1257.

Atkinson, M.R., Savageau, M.A., Myers, J.T., and Ninfa, A.J. (2003). Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli. Cell *113*, 597–607.

Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H., and Weiss, R. (2005). A synthetic multicellular system for programmed pattern formation. Nature *434*, 1130–1134.

Becker, N.A., Kahn, J.D., and Maher, L.J., III. (2005). Bacterial repression loops require enhanced DNA flexibility. J. Mol. Biol. *349*, 716–730.

Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., Kuhlman, T., and Phillips, R. (2005a). Transcriptional regulation by the numbers: applications. Curr. Opin. Genet. Dev. *15*, 125–135.

Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., and Phillips, R. (2005b). Transcriptional regulation by the numbers: models. Curr. Opin. Genet. Dev. *15*, 116–124.

Buck, M., Gallegos, M.T., Studholme, D.J., Guo, Y., and Gralla, J.D. (2000). The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. J. Bacteriol. *182*, 4129–4136.

Davidson, E.H. (2001). Genomic Regulatory Systems: Development and Evolution (Burlington, MA: Academic Press).

Davidson, E.H. (2006). The Regulatory Genome (Burlington, MA: Elsevier).

Driever, W., Thoma, G., and Nüsslein-Volhard, C. (1989). Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen. Nature *340*, 363–367.

Elowitz, M.B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. Nature *403*, 335–338.

Friedland, A.E., Lu, T.K., Wang, X., Shi, D., Church, G., and Collins, J.J. (2009). Synthetic gene networks that count. Science *324*, 1199–1202.

Garcia, H.G., and Phillips, R. (2011). Quantitative dissection of the simple repression input-output function. Proc. Natl. Acad. Sci. USA. 10.1073/pnas.1015616108.

Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in Escherichia coli. Nature *403*, 339–342.

Hervás, A.B., Canosa, I., Little, R., Dixon, R., and Santero, E. (2009). NtrC-dependent regulatory network for nitrogen assimilation in Pseudomonas putida. J. Bacteriol. *191*, 6123–6135.

Hillen, W., and Berens, C. (1994). Mechanisms underlying expression of Tn10 encoded tetracycline resistance. Annu. Rev. Microbiol. *48*, 345–369.

Hillen, W., Schollmeier, K., and Gatz, C. (1984). Control of expression of the Tn10-encoded tetracycline resistance operon. II. Interaction of RNA polymerase and TET repressor with the tet operon regulatory region. J. Mol. Biol. *172*, 185–201.

Huo, Y.X., Tian, Z.X., Rappas, M., Wen, J., Chen, Y.C., You, C.H., Zhang, X., Buck, M., Wang, Y.P., and Kolb, A. (2006). Protein-induced DNA bending clarifies the architectural organization of the sigma54-dependent glnAp2 promoter. Mol. Microbiol. *59*, 168–180.

Joung, J.K., Le, L.U., and Hochschild, A. (1993). Synergistic activation of transcription by Escherichia coli cAMP receptor protein. Proc. Natl. Acad. Sci. USA *90*, 3083–3087.

Kuhlman, T., Zhang, Z., Saier, M.H., Jr., and Hwa, T. (2007). Combinatorial transcriptional control of the lactose operon of Escherichia coli. Proc. Natl. Acad. Sci. USA *104*, 6043–6048.

Law, S.M., Bellomy, G.R., Schlax, P.J., and Record, M.T., Jr. (1993). In vivo thermodynamic analysis of repression with and without looping in lac constructs. Estimates of free and local lac repressor concentrations and of physical properties of a region of supercoiled plasmid DNA in vivo. J. Mol. Biol. *230*, 161–173.

Lederer, T., Takahashi, M., and Hillen, W. (1995). Thermodynamic analysis of tetracycline-mediated induction of Tet repressor by a quantitative methylation protection assay. Anal. Biochem. *232*, 190–196.

Lederer, T., Kintrup, M., Takahashi, M., Sum, P.E., Ellestad, G.A., and Hillen, W. (1996). Tetracycline analogs affecting binding to Tn10-Encoded Tet repressor trigger the same mechanism of induction. Biochemistry *35*, 7439–7446.

Lee, D.H., and Schleif, R.F. (1989). In vivo DNA loops in araCBAD: size limits and helical repeat. Proc. Natl. Acad. Sci. USA *86*, 476–480.

Long, T., Tu, K.C., Wang, Y., Mehta, P., Ong, N.P., Bassler, B.L., and Wingreen, N.S. (2009). Quantifying the integration of quorum-sensing signals with single-cell resolution. PLoS Biol. *7*, e68.

Magasanik, B. (1993). The regulation of nitrogen utilization in enteric bacteria. J. Cell. Biochem. *51*, 34–40.

Mukherji, S., and van Oudenaarden, A. (2009). Synthetic biology: Understanding biological design from synthetic circuits. Nature Rev. Genetics *10*, 859–871.

Müller, J., Oehler, S., and Müller-Hill, B. (1996). Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. J. Mol. Biol. *257*, 21–29.

Ninfa, A.J., and Atkinson, M.R. (2000). PII signal transduction proteins. Trends Microbiol. *8*, 172–179.

Ninfa, A.J., Reitzer, L.J., and Magasanik, B. (1987). Initiation of transcription at the bacterial glnAp2 promoter by purified E. coli components is facilitated by enhancers. Cell *50*, 1039–1046.

Orth, P., Schnappinger, D., Hillen, W., Saenger, W., and Hinrichs, W. (2000). Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. Nat. Struct. Biol. *7*, 215–219.

Phillips, R., Kondev, J., and Theriot, J. (2009). Physical Biology of the Cell (New York: Garland Science).

Qin, Y., Keenan, C., and Farrand, S.K. (2009). N- and C-terminal regions of the quorum-sensing activator TraR cooperate in interactions with the alpha and sigma-70 components of RNA polymerase. Mol. Microbiol. *74*, 330–346.

Ramos, J.L., Martínez-Bueno, M., Molina-Henares, A.J., Terán, W., Watanabe, K., Zhang, X., Gallegos, M.T., Brennan, R., and Tobes, R. (2005). The TetR family of transcriptional repressors. Microbiol. Mol. Biol. Rev. *69*, 326–356.

Rappas, M., Bose, D., and Zhang, X. (2007). Bacterial enhancer-binding proteins: unlocking sigma54-dependent gene transcription. Curr. Opin. Struct. Biol. *17*, 110–116.

Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S., and Elowitz, M.B. (2005). Gene regulation at the single-cell level. Science *307*, 1962–1965.

Schulz, A., Langowski, J., and Rippe, K. (2000). The effect of the DNA conformation on the rate of NtrC activated transcription of Escherichia coli RNA polymerase.sigma(54) holoenzyme. J. Mol. Biol. *300*, 709–725.

Su, W., Porter, S., Kustu, S., and Echols, H. (1990). DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial glnA promoter. Proc. Natl. Acad. Sci. USA *87*, 5504–5508.